(Pages : 4) **S – 1730**

Reg. No. : ....

Name : ....

Fifth Semester B.Sc. Degree Examination, December 2023

First Degree Programme under CBCSS

Statistics

Core Course VII

ST 1543 : TESTING OF HYPOTHESIS

(2018 Admission Onwards)

Time : 3 Hours Max. Marks : 80

SECTION – A

Answer **all** questions. Each question carries **1** mark.

1. Give an example of a composite hypothesis.

2. Define level of significance.

3. Define power function.

4. Define critical region.

5. What is randomized test?

6. Give the test statistics used to test the significance of population proportion.

7. Describe large sample tests.

P.T.O.

8. Give the test statistics used to test $H_0 : \mu = \mu_0$ vs $H_1 : \mu = \mu_1$ based on small sample observations from $N(\mu, \sigma^2)$, σ^2 is unknown.

9. Define runs.

10. Specify the null hypothesis in the case of median test.

(10 × 1 = 10 Marks)

SECTION – B

Answer any **eight** questions. Each question carries **2** marks.

11. Distinguish between one tailed and two tailed tests.

12. State Neyman–Pearson lemma and mention its uses in testing of hypothesis.

13. Explain uniformly most power test.

14. Describe likelyhood ratio test.

15. Describe large sample test for testing the significance of population variance.

16. Explain how will you find the degrees of freedom of chi-square test for goodness of fit.

17. Describe the test statistics used to test the equality of variances of two independent normal populations, for small samples.

18. Explain one sample sign test.

19. Describe kernal U-statistics.

20. Discuss the null hypothesis and test statistics used in the case of one sample Kalmogrou–Smirnov test.

21. Explain large sample test for testing the significance of correlation coefficient.

22. What are the disadvantages of non-parametric tests?

(8 × 2 = 16 Marks)

SECTION – C

Answer any **six** questions. Each question carries **4** marks.

23. Let P denote the probability of getting head in a single toss of a coin. For testing $H_0 : P = \frac{1}{2}$ vs $H_1 : P = \frac{3}{4}$, the coin is tossed five times and H_0 is rejected if number of heads obtained is greater than 3. Find the probabilities of type I and type II errors.

24. Obtain the most powerful test for testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$ based on a random sample of size n from $P(\lambda)$.

25. Explain large sample test for testing the equality of two population means.

26. Explain student's t test for testing the significance of mean. Nine items of a sample had the following observations.

    45, 47, 50, 52, 48, 47, 49, 53, 51.

    Do the mean of these observations differ significantly from the assumed mean 47.5

27. Explain chi-square test for testing the independence of attributes.

28. Discuss small sample test for testing the significance of correlation coefficient. A random sample of 18 pairs from a bivariate normal population showed a correlation coefficient 0.5. If this value significant from the population correlation coefficient $\rho = 0$ or not.

29. Explain the properties of likelihood ratio test.

30. Describe two sample Kolmogrov–Smirnov test.

31. Discuss Mann–Whitney Wilcoxon test.

($6 \times 4 = 24$ Marks)

SECTION – D

Answer any **two** questions. **Each** question carries **15** marks.

32. Using Neyman–Pearson lemma, find the most powerful test for testing

(a) $H_0 : \mu = \mu_0$ vs $H_1 : \mu - \mu_1$ in $N\left(\mu, \sigma_0^2\right)$, σ_0^2 is known.

(b) $H_0 : \sigma = \sigma_0$ vs $H_1 : \sigma = \sigma_1$ in $N\left(\mu_0, \sigma^2\right)$, μ_0 is known.

33. (a) Describe small sample test for testing the equality of means of two independent normal populations with common variance.

(b) Two independent samples gave the following values

Sample I	:	9	11	13	11	15	9	12	14
Sample II	:	10	12	10	14	9	8	10	

Is the difference between means significant?

34. (a) Explain large sample test for testing the equality of proportions of two independent populations.

(b) In a random sample of 500 persons in a town 200 are found to be consumers of cheese. In a sample of 400 persons from another town, 200 are found to be consumers of cheese. Examine whether the proportion of consumers of cheese in two towns are significant or not.

35. (a) Describe Wilcoxon's signed rank test.

(b) Describe two sample run test.

(c) Make a comparison between parametric and non parametric tests.

(2 × 15 = 30 Marks)

———————

Statistics

(Pages : 4) **S – 1731**

Reg. No. : ....

Name : ....

Fifth Semester B.Sc. Degree Examination, December 2023

First Degree Programme under CBCSS

Statistics

Core Course VIII

ST 1544 : SAMPLE SURVEY METHODS

(2018 Admission Onwards)

Time : 3 Hours Max. Marks : 80

Use of calculator in permitted.

SECTION – A

Answer **all** questions.

1. How many times can the same sampling unit be included in the sample, in the case of simple random sampling with replacement?

2. List out any two situations in which sampling are inevitable?

3. Define circular systematic sampling.

4. What is meant by optimum allocation?

5. Find the standard error of sample mean X' based on sample size n having variance s^2, based on a population consist of **N** items.

P.T.O.

6. Define Sampling fraction.

7. What is the condition for the ratio estimator of population mean is more efficient than estimator of population mean obtained by SRSWOR method?

8. What is the probability of any one sample of size n being drawn out of N units, in case of SRSWOR?

9. List out any two situations for the occurrences of non-sampling errors.

10. Find the variance of the sample mean in the case of SRSWR.

($10 \times 1 = 10$ Marks)

SECTION B

Answer **any eight** questions. Each question carries **2** marks.

11. Distinguish between estimate and estimator.

12. Give an example for population and sample.

13. Explain the terms : sampling unit and sampling frame.

14. A sample of 30 students is to be drawn from a population consists of 300 students belonging to two colleges of strength 200 and 100 respectively. What is the value of n_1 and n_2 if we use proportional allocation?

15. What is a sample survey?

16. What are the demerits of systematic sampling?

17. Explain the different methods of collecting primary data.

18. What do you mean by sampling and non-sampling errors? Explain various sources of non-sampling errors.

19. How will you use random number tables to select a sample in simple random sampling?

20. What is bias and explain its effects in estimators?

21. Define regression estimator of population mean.

22. What are the principles of Sample survey?

($8 \times 2 = 16$ Marks)

SECTION C

Answer **any six** questions. Each question carries **4** marks.

23. Discuss the requirements of a good questionnaire.

24. Find the unbiased estimate of population mean in stratified random sampling. Also find its variance.

25. What is linear systematic sampling? Give an example.

26. Explain the method of determining sample size in simple random sampling.

27. In SRSWOR of attributes, find an unbiased estimator for population proportion. Also find its variance.

28. Distinguish between complete enumeration and sampling.

29. In SRSWOR, show that the sample mean square is an unbiased estimate of the population mean square.

30. What does a circular systematic sampling mean? Give an example.

31. Show that in the case of sampling from finite population, the usual regression estimator is biased.

($6 \times 4 = 24$ Marks)

SECTION D

Answer **any two** questions. Each question carries **15** marks.

32. Consider a population of 6 units with values 1,2, 3,4, 5, 6.

    (a) Write down all possible samples of size 2 by SRSWOR from the population.

    (b) Verify that sample mean is an unbiased estimator of population mean.

    (c) Calculate sampling variance and show that it agrees with the formula for variance of sample mean.

    (d) Compare the efficiency of sample mean under SRSWOR and SRSWR for estimating population mean.

33. If the population consists of a linear trend, then prove that $V(\bar{y}_{st}) \leq V(y_{sys}) \leq V(\bar{y}_{srswor})$

34. Compare the efficiencies of the Neyman and proportional allocations with that of an unstratified random sample of the same size.

35. Obtain the Bias and mean squared error of ratio estimator.

(2 × 15 = 30 Marks)

———————

(Pages : 4) **S – 1729**

Reg. No. : ....

Name : ....

Fifth Semester B.Sc. Degree Examination, December 2023

First Degree Programme under CBCSS

Statistics

Core Course VI

ST 1542 : ESTIMATION

(2018 Admission Onwards)

Time : 3 Hours Max. Marks : 80

Use of statistical table and scientific calculator are permitted.

SECTION – A

Answer **all** questions. Each question carries **1** mark.

1. Define statistic. Give an example.

2. What is an estimate?

3. Define interval estimation.

4. Write the sufficient condition for the consistency of an estimator.

5. For a sample of size 18, write the 95% confidence interval for the mean of a Normal population when the population variance is unknown.

6. What is the use of Cramer-Rao inequality?

P.T.O.

7. Let T_1 and T_2 be two estimators of a parameter. When do you say that T_1 is more efficient than T_2?

8. Define moment estimator.

9. Write an example for a maximum likelihood estimator which is consistent and unbiased.

10. Define linear parametric function.

(10 × 1 = 10 Marks)

SECTION – B

Answer any **eight** questions. Each question carries **2** marks.

11. Define parameter and parameter space. Give examples.

12. Let $x_1, x_2, ..., x_n$ be a random sample with probability mass function $P(X = x) = \begin{cases} \theta, if\ x = 1 \\ 1-\theta, if\ x = 0 \end{cases}$. Show that $\frac{\sum x_i(\sum x_i - 1)}{n(n-1)}$ is an unbiased estimator of θ^2.

13. Consider a random sample of observations $x_1, x_2, .., x_n$ from a population with probability mass function $P(X = x_i) = \frac{1}{M}, x_i = 1, 2, ..., M$. Prove that $(2\overline{x} - 1)$ is a consistent estimator of M.

14. Sample variance of a random sample of 20 observations drawn from a Normal population is 6, Construct a 95% confidence interval for the population variance.

15. In a sample survey conducted at a town A, 200 residents were participated and it is found that 60% of them were in favour of tea. Determine the 99% confidence interval for the proportion of tea drinkers in town A.

16. Discuss the disadvantages of interval estimation over point estimation.

17. State Neyman factorization theorem.

18. Define (a) minimum variance unbiased estimator and (b) minimum bound estimator.

19. Let X_1, X_2 and X_3 a random sample taken from a Normal distribution with mean θ and variance 25. If $T = X_1 + X_2 + 2X_3$ and $S = 2X_1 - 3X_2 + \frac{1}{2}X_3$ are two estimates of θ, find the most efficient estimator among T and S.

20. Describe maximum likelihood estimation method.

21. Write the properties of least square estimators.

22. When do we say that a parametric function is estimable?

(8 × 2 = 16 Marks)

SECTION – C

Answer any **six** questions. Each question carries **4** marks.

23. Let $x_1, x_2, ..., x_n$ be a random sample of n observations from a population whose pdf is $f(x) = \frac{2}{\theta^2}(\theta - x)$, $0 < x < \theta$. Find the unbiased estimator of θ.

24. Consider a random sample $x_1, x_2, ..., x_n$ taken from a population with pdf $f(x) = \begin{cases} \frac{1}{b-2}, 2 < x < b \\ 0, otherwise \end{cases}$. Prove that $2(\overline{x} - 1)$ is a consistent estimator of b.

25. The sample mean and sample variance of a random sample of size 200 drawn from a population are 3.2 and 4.5 respectively. Construct the 90% and 95% confidence intervals for the population mean using the given information.

26. Obtain $(1-\alpha)100\%$ confidence interval for the difference of proportions of two Binomial populations.

27. Let X_1 and X_2 be independent and identically distributed Poisson random variables with parameter λ. Show that the Statistic $X_1 + 2X_2$ is not a sufficient estimator of λ.

28. Show that the sample mean of random sample drawn from a population with probability density function $f(x) = \frac{1}{\theta} e^{-x/\theta}$, $0 < x < \infty$ is the MVB estimator of θ.

29. Estimate the parameter θ of the population with Probability density function $f(x) = e^{-(x-\theta)}$, $0 < x < \infty$ using the method of moments.

30. Find the maximum likelihood estimate of the parameter p for the distribution with probability mass function $P(X = x) = {}^{n}C_x p^x (1-p)^{n-x}$, $0 < p < 1$, $x = 0,1,2,\ldots,n$. Also check the estimate is unbiased or not?

31. Let Y_1, Y_2, Y_3 be uncorrelated observations with common variance σ^2 and expectations given by $E(Y_1) = \beta_0 + \beta_2$, $E(Y_2) = \beta_1 + \beta_2, E(Y_3) = \beta_2 + \beta_3$, where β_i's are unknown parameters. Derive the criterion for estimability and check whether $\beta_1 + \beta_2 + \beta_3$ and $\beta_1 + \beta_2 - \beta_3$ are estimable or not?

(6 × 4 = 24 Marks)

SECTION – D

Answer any **two** questions. **Each** question carries **15** marks.

32. Prove that for the Cauchy distribution with pdf $f(x) = \frac{1}{\pi\left[1 + (x-\mu)^2\right]}, -\infty < x < \infty$ sample mean is not but the sample median is the consistent estimator of μ.

33. Consider a Normal population with mean μ and variance σ^2. Obtain $(1-\alpha)100\%$ confidence interval for (a) population mean μ when σ is known and (b) population variance σ^2.

34. Let $x_1, x_2, \ldots, x_n$ be a random sample of n observations from a population whose pdf is $f(x) = \frac{1}{\Gamma(m)\theta^m} e^{-x/\theta} x^{m-}, m > 0, x > 0$. Find the maximum likelihood estimator of θ when (a) m is known (b) $m = 1$ and (c) $m = 2$.

35. Establish Gauss Markov theorem.

(2 × 15 = 30 Marks)

(Pages : 4) **S – 1728**

Reg. No. : ....

Name : ....

Fifth Semester B.Sc. Degree Examination, December 2023

First Degree Programme under CBCSS

Statistics

Core Course V

ST 1541 : LIMIT THEOREMS AND SAMPLING DISTRIBUTIONS

(2018 Admission Onwards)

Time : 3 Hours Max. Marks : 80

SECTION – A

Answer **all** questions. Each question carries **1** mark.

1. If $A \subseteq B$ prove $P(A) \leq P(B)$, using axioms of Probability.

2. Define probability space.

3. Suppose $\{X_n\}$ and $\{Y_n\}$ converges to X and Y respectively in probability. What can you say about the convergence of $\{X_n.Y_n\}$?

4. Define convergence in distribution of a sequence of random variables $\{X_n\}$.

5. If $X \sim \chi^2_{(3)}$ and $Y \sim \chi^2_{(5)}$ are two independent Chi-square random variables, what is the mean of X+Y?

6. If X is a random variable with Student's t distribution with n degrees of freedom, what is the distribution of X^2?

7. Define non-central F distribution.

8. If $X \sim F(m,n)$, what is the distribution of 1/X?

9. If $X_1, X_2, ..., X_n$ be a random sample from $U(0,\theta)$, find the distribution function of $Max(X_1, X_2, ... X_n)$.

10. Define order statistics.

(10 × 1 = 10 Marks)

SECTION – B

Answer any **eight** questions. Each question carries **2** marks.

11. Write down the expressions for limit supremum and limit infimum for a sequence of events $\{A_n\}$.

12. State Borel - Cantelli lemma.

13. Suppose $\{X_n\}$ is a sequence of random variables with probability mass function $P(X_n = 1) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$. Examine the convergence in probability of $\{X_n\}$.

14. Let $\{X_n\}$ be a sequence of random variables with distribution function $F_{X_n}(x) = 1 - (1 - \frac{1}{n})^{n.x}; x > 0$ and 0 otherwise. Show that $\{X_n\}$ converges in distribution to an exponential distribution with unit mean.

15. A random variable X has mean 5 and variance 3. Find the least value of $P(|X - 5| < 7.5)$ using Chebyshev's inequality.

16. A random sample of size 64 are drawn from a population with mean 32 and standard deviation 5. Find the mean and standard deviation of the sample mean $\overline{X}$.

17. If (X_1, X_2, X_3) is a random sample of size 3 from a standard normal population $N(0,1)$, what is the sampling distribution of $U = \frac{\sqrt{2}X_3}{\sqrt{X_1^2 + X_2^2}}$.

18. What are the assumptions on Central Limit theorem?

19. If the sample values are 1,3,5,6,9, find the standard error of the sample mean.

20. List the advantages of Chebychev's inequality.

21. Let $(X_1, X_2, ... X_n)$ be a random sample of size n from a distribution with density function f(x) and distribution function F(x). Write down the distribution function and probability density function of the n^{th} order statistic, $X_{(n)}$.

22. What is meant by weak law of large numbers?

(8 × 2 = 16 Marks)

SECTION – C

Answer any **six** questions. Each question carries **4** marks.

23. State and prove Bernoulli's law of large numbers.

24. If $A_1, A_2, ..., A_n ..$ is a sequence of events in sample space S such that $A_1 \subseteq A_2 \subseteq ... \subseteq A_n ..$ then prove that $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \to \infty} P(A_n)$.

25. Let $\{X_k\}$ be a sequence of independent random variables with values $-2^k, 0$ and 2^k and probabilities $P(X_k = \pm 2^k) = 2^{-(2k+1)}$; $P(X_k = 0) = 1 - 2^{-2k}$. Examine whether weak law of large numbers holds for the sequence.

26. Let the probability density function of a random variable X be $f(x) = 1;\ 0 < x < 1$. What is the lower bound of $P\left(\left|X - \frac{1}{2}\right| \le 2\sqrt{\frac{1}{12}}\right)$ when one uses the Chebyshev's inequality?

27. Let $\overline{X}$ be the sample mean of a random sample of size 50 from a normal population with mean 112 and standard deviation 40, Find (a) $P(110 < \overline{X} < 114)$ (b) $P(\overline{X} > 113)$.

28. State and prove additive property of Chi square distribution.

29. Let (X_1, X_2) be a random sample from a distribution with density function $f(x) = e^{-x}; x > 0$. Find the density function of $Y = \min(X_1, X_2)$.

30. Let $X_1, X_2, .., X_n$ be a random sample from a uniform distribution over (0,1). Find the probability density function of r^{th} order statistic $X_{(r)}$.

31. Let $X_1, X_2, .., X_n$ be n independent variates, X_i having Geometric distribution with parameter p_i, i.e., $P(X_i = x_i) = q_i^{x_i - 1} p_i$, $q_i = 1 - p_i$, $x_i = 1,2,3, ...$ Show that $X_{(1)}$ is distributed geometrically with parameter $(1 - q_1 q_2 ... q_n)$

(6 × 4 = 24 Marks)

SECTION – D

Answer any **two** questions. **Each** question carries **15** marks.

32. (a) If X is a continuous random variable with mean μ and variance σ^2, establish Chebyshev's inequality.

    (b) If X is a random variable with $E(X) = 3$ and $E(X^2) = 13$, use Chebyshev's inequality to determine the lower bound for the probability P(-2<X<8).

33. (a) State and prove Lindberg-Levy form of central limit theorem.

    (b) If $X_1, X_2, ..., X_n$ is a sequence of Bernoulli random variables with probability success p. write down the central limit theorem result.

34. (a) Define t, χ^2 and F statistics and give relationship between each of them.

    (b) Obtain r^{th} arbitrary moment μ'_r of F distribution with (m, n) degrees of freedom.

35. (a) Derive the sampling distribution of means of samples chosen from a normal population.

    (b) A population is known to follow normal distribution with mean 2 and S.D. 3. Find the probability that the mean of a sample size 16 taken from this population will be greater than 2.5.

(2 × 15 = 30 Marks)

(Pages : 4) **S – 1733**

Reg. No. : ....

Name : ....

Fifth Semester B.Sc. Degree Examination, December 2023

First Degree Programme Under CBCSS

Statistics

Open Course

ST 1551.4 — OFFICIAL STATISTICS

(2018 Admission Onwards)

Time : 3 Hours Max. Marks : 80

(Use of calculator is allowed)

SECTION – A

Answer **all** questions. Each question carries **1** mark.

1. Expand CSO.

2. Who is the founder of the Indian Statistical Institute?

3. What is the sex ratio of India as per the population census 2011?

4. Define crude death rate.

5. Define secular trend.

6. Define cyclic variations.

7. Give any two uses of time series.

P.T.O.

8. Define index number.

9. Which index number is considered as ideal for index number constructions?

10. Give any two limitations of index numbers.

(10 × 1 = 10 Marks)

SECTION – B

Answer any **eight** questions. Each question carries **2** marks.

11. Explain the importance of statistical system in the state.

12. Give any four publications of CSO.

13. What do you mean by the population pyramld?

14. Define force of mortality.

15. What is time series? Mention its chief components.

16. Explain the mixed models in time series.

17. Define specific fertility rate.

18. Define stationary population.

19. Explain consumer price index number.

20. Explain the simple aggregate method of construction of index numbers.

21. Give the computation formula for Fisher's price index.

22. Explain the seasonal variations in time series.

(8 × 2 = 16 Marks)

SECTION – C

Answer any **six** questions. Each question carries **4** marks.

23. Explain the various activities of CSO.

24. Explain the various columns of a life table.

25. Describe the standardized death rates.

26. Explain the gross reproduction rate.

27. Explain the method of semi average for the measurement of trend.

28. Explain the fitting of linear trend $y = a + bt$, where a and b are constants and t is time.

29. Explain the moving average method of the measurement of trend.

30. Explain factor reversal test of index numbers and give the index number which satisfies this test.

31. What are the major problems involved in the constructions of an index number?

(6 × 4 = 24 Marks)

SECTION – D

Answer any **two** questions. Each question carries **15** marks.

32. What are the various divisions of NSSO? Explain the responsibilities of these divisions.

33. Fill in the blanks in a portion of life table which are marked a query?

Age in years:	l_x	d_x	p_x	q_x	L_x	T_x	e_x^0
4	95000	500	?	?	?	4850300	?
5	?	400	?	?	?	?	?

34. From the following data calculate price index numbers For 2010 with 2000 as base year by (a) Laspeyre's method, (b) Paasche's method, and (c) Fisher's Ideal method.

Commodities	2000		2010	
	Price	Quantity	Price	Quantity
A	500	100	900	150
B	320	80	500	100
C	150	60	360	72
D	360	30	297	33

35. The sales of a company in lakhs of rupees for the years 1994 – 2001 are given below.

Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Sales:	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

Fit the linear trend equation and compute the trend values for 1991 and 1996.

($2 \times 15 = 30$ Marks)